

THE IMPACT OF TRAINING ON PRODUCTIVITY AND WAGES: FIRM LEVEL EVIDENCE*

Jozef Konings

Vives, Department of Economics, KU Leuven, Belgium

and University of Ljubljana, Slovenia

Stijn Vanormelingen

KU Leuven and HU Brussels, Belgium

April 9, 2014

*We would like to thank the editor Gordon Hanson and two anonymous referees for their helpful comments. Moreover we are grateful to Filip Abraham, Italo Colantone, Anneleen Forrier, Lisa George, Johannes Van Biesebroeck, Patrick Van Cayseele, Bart Cockx, Marton Csillag, Jan De Loecker, Garth Frazer, Guido Friebel, Steve Pischke, Bee Roberts and Vitor Gaspar for useful comments and suggestions on earlier drafts of the paper. This paper has benefited from various presentations at conferences and seminars.

Abstract

This paper uses firm level panel data of on-the-job training to estimate its impact on productivity and wages. To this end, we apply and extend the control function approach for estimating production functions, which allows us to correct for the endogeneity of input factors and training. We find that the productivity premium of a trained worker is substantially higher compared to the wage premium. Our results are consistent with recent theories that explain work related training by imperfect competition in the labor market.

JEL codes: J24, J31, D24

Keywords: Training, production functions, human capital.

1 Introduction

The accumulation of human capital plays an important role in explaining economic performance and long-term growth (Lucas 1988). Mostly the focus lies on skill acquisition through the general education system. However, on-the-job training plays a crucial role as well because it can not only maintain, but also improve human capital of the workforce. While there exists a vast literature estimating the returns to training, which focused mainly on the impact on wages¹, there are only a few papers that also analyzed the impact of training on productivity.² Moreover, the focus in these papers is either on the impact on wages or on the productivity premium of training. In contrast, this paper analyzes the impact of on-the-job training on both wages and productivity, which matters for understanding the economic mechanisms behind training.

The theoretical foundations of on-the-job training have originally been formalized by Becker (1964) who made a distinction between general and specific training. Under

¹Using employee level datasets, large and significant effects of work related training on wages are usually found ranging between 1.1% and 16.6%. Notable examples include Altonji and Spletzer (1991), Lynch (1992), Loewenstein and Spletzer (1999), Parent (1999) and Frazis and Loewenstein (2005) for the U.S and for Europe, Booth (1991), Pischke (2001), Booth, Francesconi and Zoega (2003) and Booth and Bryan (2005) among others. For an overview see Bassanini et al. (2007).

²These papers report mixed results but are only based on limited samples (Bartel 1995, Black and Lynch 2001, Zwick 2006). Moreover, they do not analyze wage and productivity premia together. An exception is Dearden, Reed and Van Reenen (2006), analyzing the link between training, wages and productivity at the sector level using a panel of British industries.

perfect competition, firms will not pay for general training of their workers as they can leave the firm searching for better paid work, which compensates them for the increased productivity acquired through general training. Hence, the worker is the sole recipient of general training benefits and will also bear the costs of it. Yet, in a series of papers Acemoglu and Pischke (1998, 1999a, 1999b) argue that a substantial amount of training is paid for by firms and is still general in nature. They show that a necessary condition for firms to pay for general training is a compressed wage structure, caused by imperfections in the labor market such as monopsony. With a compressed wage structure, training increases the marginal product of labor more than the wage, which creates incentives for the firm to invest in general training.

Our paper contributes to the literature along various dimensions. First, we make use of a large firm level longitudinal data set which contains information on measures of training, such as the proportion of workers that received training, the number of hours they were trained and the cost of training. This data allows us to measure the impact of training on both wages and productivity at the firm level. By focusing on firm level data we are able to avoid possible aggregation biases and hence capture the effects of training more precisely.³ Second, the analysis at the firm level and the panel structure of the data allows us to control for the endogeneity of training. To this end, we estimate production functions applying recent control function approaches taking into account training decisions. In addition, the production function estimates provide us with a measure of unobserved worker ability which we include in the wage equation to retrieve a consistent estimate for the impact of training on wages as in Frazer (2001). Third,

³Because we use firm level data on training, we do not capture spillovers in human capital across firms as opposed to Dearden et al. (2006).

our data allows us to explore how the impact of training on wages and productivity is affected by worker heterogeneity related to the type of worker contracts, human capital and gender.

We find that an increase in the share of trained workers by 10 percentage points is associated with 1.7 percent to 3.2 percent higher productivity, depending on the specification. However, consistent with the theoretical insights about wage compression and training, this increase in productivity is not entirely offset by a similar increase in wages. The average wage per worker only increases by 1.0 to 1.7 percent in response to the same increase in training.

In the next two sections we develop our empirical framework and estimation strategy. Section 4 introduces the data. We report our results in Section 5, including a battery of robustness checks both in terms of the empirical specification and estimation method. Section 6 distinguishes between firm specific and general training and Section 7 concludes.

2 Empirical Framework

We infer the impact of on-the-job training on both wages and productivity by applying a framework similar to Hellerstein, Neumark and Troske (1999), which has been commonly used to compare returns to characteristics of the labor force such as gender, race and human capital on both wages and productivity. The idea is essentially to estimate both a production function and wage equation to infer productivity and wage premia for the different labor force characteristics. In competitive labor markets, the wage premium associated with each worker characteristic should equal the corresponding productivity premium. As it is not possible to observe the individual contributions of workers to output, some aggregation of employee and output data is necessary, as reported in firm

or plant level data. We depart from the standard framework of Hellerstein et al. (1999) and allow for continuous worker characteristics (Frazer 2001, Van Biesebroeck 2011). We next outline our empirical approach to infer the impact on productivity and on wages. A more detailed description is given in Appendix A.

2.1 Impact of training on productivity

The output of a firm i in period t is a function of capital and a labor quality aggregate used by the firm in period t . As is common in the literature, we assume that this function takes the Cobb-Douglas form:

$$Y_{it} = \widehat{L}_{it}^{\beta_l} K_{it}^{\beta_k} \exp(q_{it}) \exp(\varepsilon_{it}) \quad (1)$$

where Y_{it} represents value added, \widehat{L}_{it} is aggregate effective labor input, K_{it} is capital and q_{it} represents technical efficiency shifting the production function. Suppose for the moment that workers can be distinguished according to their education and training level. If these characteristics enter the effective labor input as in a Mincer (1974) wage equation, the labor aggregate at the firm level can be written as $\ln \widehat{L}_{it} = \ln L_{it} + \beta_T \overline{T}_{it} + \beta_S \overline{S}_{it} + Z_{it}$ (cf. Appendix A). Here, \overline{T}_{it} represents the average training intensity of the workforce employed in firm i during period t . Likewise \overline{S}_{it} represents the average schooling level of the workforce and Z_{it} is unobserved labor quality. The parameters β_T and β_S are the productivity premia associated with training and schooling respectively. The production function can subsequently be written as follows:⁴

⁴Throughout the rest of the paper, lower case letters represent variables expressed in logarithms.

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_l \beta_T \bar{T}_{it} + \beta_l \beta_S \bar{S}_{it} + \omega_{it} + \varepsilon_{it} \quad (2)$$

When training and schooling are discrete variables, the average schooling and training levels are just equal to the proportions of trained and schooled workers, L_T/L and L_S/L , in the labor force and the estimation equation is exactly the same as the one derived by Hellerstein et al. (1999). Unobserved productivity ω_{it} includes both technological progress and unobserved labor quality. Our main parameter of interest is β_T , which measures how the labor aggregate varies with training intensity ($\partial \ln \hat{L} / \partial \bar{T} = \beta_T$) and reflects the impact of training on the marginal product of a worker. If training intensity is defined as a discrete characteristic the parameter reflects the productivity premium of a trained worker compared to an untrained worker. The impact of training on output depends as well on the importance of labor in the production technology, i.e. $\partial y / \partial \bar{T} = \beta_l \beta_T$ which represents the percentage changes in output in response to variations in the training intensity of the workforce.

2.2 Impact of training on wages

Applying a similar derivation as for the labor aggregate in the production function, Appendix A shows how the logarithm of the average wage, \bar{w}_{it} , paid by firm i in period t can be written as:

$$\bar{w}_{it} = w_0 + \alpha_T \bar{T}_{it} + \alpha_S \bar{S}_{it} + \alpha_0 Z_{it} \quad (3)$$

where again \bar{T}_{it} and \bar{S}_{it} represent the average training and schooling level respectively. Unobserved labor quality is represented by Z_{it} . Similar to Hellerstein et al. (1999) we add industry and year effects to the estimation equation as well as observed firm characteristics X_{it} such as the capital-labor ratio and an additive i.i.d. error term ε_i . The

equation that will be estimated is the following:

$$\bar{w}_{it} = w_0 + \alpha_T \bar{T}_{it} + \alpha_S \bar{S}_{it} + X_{it}\beta + \alpha_0 Z_{it} + \varepsilon_{it} \quad (4)$$

The parameters α_T and α_S measure how wages change in response to training and schooling respectively and can be compared with the impact of these human capital measures on the marginal product of workers, β_T and β_S . When worker characteristics are discrete, the Hellerstein et al. (1999) framework leads to an identical estimation equation as in (4). Applying OLS to the above equations (2) and (4) could result in biased estimates of the wage premia since the human capital variables are likely to be correlated with unobserved labor quality. We will show in the next section how we will obtain consistent estimates for the parameters.

3 Estimation strategy

To identify the differential impact of training on both wages and productivity, we need to consistently estimate the coefficients of both the production function and the wage equation. First we describe how we estimate the production function, next we show how these estimates help us to identify the training impact. Recall the production function derived in the previous section and assume for simplicity that workers are only distinguished by one discrete observable characteristic, namely training:

$$y_{it} = \beta_0 + \beta_k k_{it} + \beta_l l_{it} + \beta_{tr} \frac{L_{T,it}}{L_{it}} + \omega_{it} + \varepsilon_{it} \quad (5)$$

where β_{tr} is defined as $\beta_{tr} \equiv \beta_l \beta_T$. As is well known since the work by Marschak and Andrews (1944), the input choices of a profit maximizing firm are likely to be correlated with unobserved productivity ω_{it} . To control for this we apply the estimation proce-

cedure proposed by Akerberg, Caves and Frazer (2006) using the insight that optimal input choices hold information about unobserved productivity. More precisely we rely on material demand,

$$m_{it} = f_t \left(\omega_{it}, l_{it}, \frac{L_{T,it}}{L_{it}}, k_{it} \right) \quad (6)$$

where we assume labor input and training to be set before the choice of material input. If material demand – conditional on labor, capital and training – is monotonically increasing in productivity, the function can be inverted and productivity can be expressed as a function of observables. Substituting this inverted material demand function in the production function, results in the first stage regression equation:

$$y_{it} = \beta_l l_{it} + \beta_{tr} \frac{L_{T,it}}{L_{it}} + \beta_k k_{it} + f_t^{-1} \left(m_{it}, l_{it}, \frac{L_{T,it}}{L_{it}}, k_{it} \right) + \varepsilon_{it} \quad (7)$$

We run regression equation (7) using a polynomial in materials, labor, capital and training to proxy the inverse material input function $f_t^{-1}(\cdot)$ and retrieve an estimate for expected output $\phi_{it} = \beta_l l_{it} + \beta_{tr} (L_{T,it}/L_{it}) + \beta_k k_{it} + f_t^{-1}(\cdot)$. The input coefficients, β_l , β_k and β_{tr} will be identified in the second stage. An important advantage of this procedure, given our research question and the peculiarities of the Belgian labor market, is that it is consistent with labor choices having dynamic implications due to for example hiring, firing or training costs. Although labor and capital will depend on lagged labor in this case, optimal material demand m_{it} will only be a function of l_{it} , $L_{T,it}/L_{it}$ and k_{it} as it is only relevant for production in period t . However, there cannot exist unobservables that directly affect material demand since they would make the inversion of the material demand function invalid. Moreover, the identification strategy rests on the assumption that materials are chosen at the same time production takes place. Given the large heterogeneity across sectors in the materials used, we will perform a robustness check on a subsample of sectors that are more likely to purchase readily available inputs.

The second stage of the estimation procedure serves to identify all the input coefficients. As is standard in the literature we assume ω_{it} to follow a first order Markov process and we can write $\omega_{it} = g_t(\omega_{it-1}) + \xi_{it}$ where ξ_{it} represents a productivity shock unexpected in period $t - 1$. After the first stage we can compute productivity ω_{it} for every candidate vector of input coefficients $\beta = (\beta_l, \beta_k, \beta_{tr})$ and non-parametrically regressing $\omega_{it}(\beta)$ on $\omega_{it-1}(\beta)$ allows us to recover the productivity shock $\xi_{it}(\beta)$. We can now use our timing assumptions to form the moment conditions used to identify the input coefficients. First, we keep the standard assumption about capital accumulation, namely that investment decided in period t only enters the capital stock in period $t + 1$. Consequently the capital stock in period t will be uncorrelated with the unexpected productivity shock in period t . Moreover, we assume that labor input and the amount of training do not depend on the innovation in productivity. For the labor coefficient, this is a more strict assumption than usually applied but can be justified by the substantial labor adjustment costs in Belgium.⁵ We report as well results where we relax this assumption and use lagged values of labor instead of current labor to construct the moment conditions. Concerning the training variable, several human resource managers confirmed that the amount of training provided to workers is mostly decided one year in advance when making up the budget for the following year, which makes the amount of training independent from the innovation in productivity, $\xi_{it}(\beta)$. Consequently, the moment conditions to identify the

⁵For example, the OECD Employment Protection Legislation Index shows the importance of substantial adjustment costs for a large number of countries among which Belgium has one of the highest scores, especially for the notice and severance pay for individual dismissals, legislation concerning collective dismissals and temporary employment (OECD 2007).

input coefficients in the second stage are

$$E \left[\xi_{it} \begin{pmatrix} k_{it} \\ l_{it} \\ L_{T,it}/L_{it} \end{pmatrix} \right] = 0 \quad (8)$$

and we bring the sample analogue of these moment conditions as close as possible to zero by adjusting the input coefficients.

Estimating the wage equation (4) by OLS could result as well in biased estimates of the wage premia since training is likely to be correlated with unobserved labor quality Z_{it} . To control for unobserved labor quality we can use our estimate for ω_{it} from the production function. If the main component of ω_{it} after controlling for industry and year specific effects, is labor quality (so $\omega_{it} = \omega_j + \omega_t + Z_{it}$), then adding estimated total factor productivity to the wage equation will result in the following equation to be estimated:

$$\bar{w}_{it} = w_0 + \alpha_T \bar{T}_{it} + X_{it} \beta + \alpha_0^* Z_{it} + \omega_j + \omega_t + \varepsilon_{it} \quad (9)$$

and the estimation of the equation renders consistent estimates of the wage premia. If the estimate for total factor productivity from the production function includes as well other factors than labor quality, ω_{it} imperfectly controls for labor quality in the wage equation. Consequently, our measures for the wage premia could still be upward biased. However, note that this bias works against our main conclusion that the productivity premium exceeds the wage premium.⁶

We test each time for the equality of the productivity and wage premia. Only under the joint assumptions of general training and perfect competition in the labor market, the training coefficient in the wage equation will be equal to productivity premium obtained

⁶More details on the estimation strategy are provided in Appendix A.

from the production function. If equality is rejected and training is – as we argue – general in nature, the underlying assumption of competitive spot labor markets can be discarded. Consequently, the coefficient on the training variable in the wage equation cannot be interpreted any more as the wage premium of an individual trained worker, reflecting its productivity premium. For example, with monopsonistic labor markets, the coefficient on training is likely to be a mix of the training premium at the individual level and parameters from the labor supply process. The coefficient however can still be interpreted in a reduced form way as the increase in the wage bill in response to an increase in training.

All regressions include year and industry dummies. Industry dummies are at the NACE 2 digit level for estimations on the whole sample and at the NACE 4 digit level for regressions at the sector level. Standard errors for all coefficients in both the production function and wage equation are obtained by using a block bootstrap procedure with 500 replications.

4 Data Description

Data is obtained from the Belfirst database. This database, commercialized by Bureau Van Dijck, includes the income statements of all Belgian incorporated firms. We obtained an unbalanced panel for the period 1997-2006 of both manufacturing and non-manufacturing firms with at least one worker. We select a number of key variables needed for estimation of the production function and wage equation such as value added, number of employees (in full time equivalents), labor costs, material costs and the capital stock.

In addition, Belgian firms are required to report information about formal training⁷ they provide to their employees. In particular, they have to report the number of employees that followed some kind of formal training as well as the hours spent on this training and the training costs. This allows us to obtain a firm-level measure of training for more than 135,000 Belgian firms active in manufacturing and non-manufacturing sectors. However, only a fraction of these firms have to report material costs, which we will need in our empirical strategy.⁸ A more elaborate discussion of the dataset is included in Appendix B.

Table 1 provides some summary statistics of the full dataset as well as of the restricted sample of firms reporting material costs. A Belgian firm active in the private sector employs on average 21.6 employees, generates around 1.3 million euros value added per year and has an average labor cost of around 35,400 euro. Manufacturing firms are on average larger compared to non-manufacturing firms.⁹ The average proportion of trained workers is equal to 3.2%, mainly due to the low number of firms providing training to their employees. In firms that train their workers in a given period, around 50%

⁷Formal training excludes training that takes place at the work floor or self study. The training has to take place at a separate training room or work floor especially developed for training activities. Training can take place inside or outside the firm.

⁸Only large firms in Belgium have to submit a full version of the annual report. Smaller firms only have to submit a shorter version which does not include material costs. Firms are defined to be large if they have on average more than 50 employees, realize a turnover of more than 7.3 million euro or report a total value of assets of more than 3.65 million euro.

⁹Manufacturing firms are firms active in NACE Rev. 1.1 sectors 15 to 36. The other sectors are pooled together as non-manufacturing sectors.

of the employees benefit from this training which lasts approximately one work week, namely 39.1 hours and costs 1,414 € to the firm. The training duration and costs are somewhat larger in the manufacturing sector compared to the non-manufacturing sector. We report as well the summary statistics for the subsample of firms reporting material costs as we need to observe material costs to control for the endogeneity of inputs. The subsample consists of typically larger firms which are more likely to provide training to their employees. The costs and duration of training however, are approximately the same as in the full sample. The data appendix shows more summary statistics on sector level heterogeneity in training.

5 Results

This section presents the results of the empirical analysis. First, we estimate productivity and wage premia for all sectors pooled together. We show moreover that these findings are robust to distinguishing between blue and white collar workers and hold at the sector level as well. Next we measure training as a continuous variable. Finally we perform a number of further robustness checks.

5.1 General Results

Baseline Results

Table 2 shows the results of estimating equations (5) and (9) for all firms active in all sectors pooled together and for manufacturing and non-manufacturing separately. The first column for each subsample (total, manufacturing and non-manufacturing) reports the estimation results for the full sample by applying ordinary least squares (OLS1).

Our estimation strategy to control for the endogeneity of inputs requires a measure for materials input, which is only observed for a subsample of firms. To adequately assess the performance of our estimation strategy, we report in the second column results for least squares estimation (OLS2) on this subset of firms. The third column presents the coefficient estimates obtained by following the estimation strategy outlined in Section 3. The estimates reported in column (1) show that training has a statistically significant and positive effect on productivity. The coefficients imply that a 10 percentage point increase in the proportion of trained workers is associated with 4.6% rise in value added. Turning to the subset of firms that report material costs, the coefficient on training in the production function drops somewhat to 0.315 but remains highly significant.¹⁰ Controlling for the endogeneity of inputs (and training) causes the training coefficient to drop further to 0.243 as shown in column (3). The estimates imply that value added increases by 2.4% in response to an increase of 10 percentage points of the share of trained workers such that even after controlling for the possible endogeneity of training, there remains a substantially large impact of training on productivity. Note that the results imply that on average the marginal product of a trained worker is around 32% ($\beta_T = .243/.764$) higher than the marginal product of an untrained worker. One has to bear in mind that this is an estimate for the average effect of training on the marginal product of all workers pooled over all sectors.¹¹ The results for manufacturing industries and non-manufacturing separately are comparable, although we find a slightly stronger

¹⁰The decrease in the estimated training premium is due to large firms being more productive compared to small firms and being more likely to train their workers as well (cf. Appendix B).

¹¹Moreover, when there exist spillover effects to untrained workers within a firm, our measure includes these effects and the direct impact of training will be lower.

impact of training in non-manufacturing sectors.

For the wage equation, we estimate as well three specifications. First, the log wage is regressed on the share of trained workers together with year and sector dummies (OLS1). Second, this exercise is repeated, but the sample is now restricted to firms included in the productivity estimation sample where we control for the endogeneity of inputs. As a result, the coefficient on training drops from 0.438 to 0.200. Note that although the OLS point estimates for the productivity and wage premia are likely to be biased upwards due to unobserved labor quality, the bias would affect the estimated training coefficients to a similar extent in both the production function and wage equation. Consequently, the test for equality of the premia, discussed below, remains informative. In the third specification, we add controls in the wage equation. In particular, we add the capital-labor ratio and total factor productivity as control variables, as discussed in Section 3. The coefficient on training further drops to 0.167, implying the wage premium for a trained worker in the Belgian private sector to be equal to 17%.

Table 2 shows that the productivity premium of training is larger than the wage premium and the difference is statistically significant as indicated by the Wald test of the equality of α_T and β_T .¹² The productivity premium for a trained worker is almost twice as high as his wage premium for all sectors pooled together in column (3). With a χ^2 -square statistic of 128.2, the null of equal coefficients can be rejected at any conventional significance level. The same is true for the manufacturing sector and non-manufacturing sector separately with Chi-square values of 14.1 and 113.0 respectively. The finding that

¹²To retrieve an estimate for β_T , we divide the coefficient on the share of trained workers by the labor coefficient. Consequently, the null is $(\beta_{tr}/\beta_l) - \alpha_T = 0$. This non-linear hypothesis is tested by using a Wald test.

the impact of training on productivity is higher than the impact on wages, gives support to the Acemoglu and Pischke (1998, 1999a,b) model that explains why firms invest in the general training of their employees. A necessary condition is that productivity of employees increases more than their wages in response to training.¹³ In the last three columns of Table 2, we relax the assumption that contemporaneous labor is uncorrelated with the unexpected shock in productivity and use lags of the labor variable as instruments for each subsample. As can be seen from the table, the results remain qualitatively the same.

Blue and White Collar Workers

There could be concerns that our methodology does not fully control for worker heterogeneity and our training coefficient is driven by differences in the marginal product between different types of workers. As such, the differential impact of training on wages and productivity could reflect wage-productivity gaps of these characteristics. One important dimension of worker heterogeneity is the distinction between blue collar and white collar workers. These different contract types can pick up differences in education levels across employees. Moreover employment protection in Belgium differs between blue and white collar workers.¹⁴ We bring in this extra information in our empirical framework by

¹³Note that Becker (1964) also allows for the possibility that firms pay (part of) the training costs. For this to be the case, the training needs to be firm specific in nature.

We will turn back to this issue in the last subsection.

¹⁴For the whole sample, around 52% of the workforce is blue-collar, 44% white collar and 1.4% management. In the manufacturing sector the shares are respectively 66%, 31% and 1.6% and in the services sectors respectively 45%, 51% and 1.3%. The percentages do not sum up to 100% because some of the workers have an undefined contract and can not be classified.

including two different labor aggregates in the Cobb-Douglas production function, one for blue collar workers and one for white collar workers.

$$Y_{it} = A_{it} K_{it}^{\beta_k} \widehat{L}_{B_{it}}^{\beta_b} \widehat{L}_{W_{it}}^{\beta_w} \quad (10)$$

with \widehat{L}_B and \widehat{L}_W the labor aggregate for blue and white collar workers respectively.

Assuming the share of trained workers is constant across the different types of contract, the equation that we seek to estimate becomes:

$$y_{it} = \beta_k k_{it} + \beta_b (l_B)_{it} + \beta_w (l_W)_{it} + (\beta_b \beta_{TB} + \beta_w \beta_{TW}) \frac{L_{T,it}}{L_{it}} + \omega_{it} + \eta_{it} \quad (11)$$

where β_{TB} and β_{TW} represent the productivity premium of a trained blue collar worker and the productivity premium of a trained white collar worker respectively. These premia are measured relative to an untrained worker with the same type of contract. The drawback of this specification is that we have to exclude all observations without blue or white collar workers. We do not include managers as a separate category as only a small fraction of firms reports information on the number of managers, for those that do, we simply add them to the number of white-collar workers. For the same reason we choose not to relax the assumption of perfect substitutability between trained and untrained employees. While it is theoretically possible to allow for imperfect substitution between trained and untrained employees, we would be forced to drop most of the observations since a large fraction of the firms does not provide training.

We estimate equation (11) applying our estimation strategy outlined in section 3, but we use a different timing assumption. In Belgium, white collar workers are well protected against dismissal while blue collar workers face less strict employment protection legislation. Consequently we treat here blue collar workers as an input that is adjusted in reaction to unexpected productivity shocks. To control for this we use blue collar

workers lagged one period as an instrument instead of the contemporaneous stock of blue collar workers. Results are reported in Table 3. If we assume the impact of training on productivity is the same for blue and white collar workers, the estimated coefficient implies a productivity premium of 22.8% and a wage premium of 13.0%. These estimates are slightly lower compared to the baseline results in Table 2, but still show substantial returns to training both in terms of productivity as in terms of wages. The difference between the two premia remains highly significant with a χ^2 -square statistic of 16.0 such that we can reject the equality of the productivity and wage premia at each conventional significance level. In section 5.4, we perform some additional robustness checks related to worker heterogeneity.

The estimated wage premium falls within the range, albeit at the higher end, of wage premia found in other studies. These studies mostly use employee level data and premia go from 4 to 16 percent. Concerning the impact on productivity, Dearden et al. (2006), using industry level data, find that an increase of 1% in their training measure is associated with an increase in value added per hour of about 0.6% – implying productivity premia of over 60% – and an increase in the average wage of about 0.3%, substantially larger than our estimates. However, note that the median training duration in their sample is around 2 weeks, twice as long as in the current sample, such that the productivity and wage premium of an hour of training are more comparable. The remaining difference could be due to the different level of aggregation.

Sector Heterogeneity

So far, we assumed the same production technologies as well as training effects across the different sectors. In contrast to previous studies we can relax this assumption and allow

for sector heterogeneity in our coefficient estimates. In Tables 4 and 5 we report results for each NACE 2 digit sector separately. For brevity, we only report the effect of training on wages and worker’s marginal product together with the χ^2 statistic and p -value of the Wald test for testing the equality of the productivity and wage premia. The coefficients on the other regressors are reported in Appendix C. For the majority of sectors, both the labor and training coefficients go down in the production function and wage equation when controlling for their possible endogeneity. The unweighted average for the coefficient – controlling for the endogeneity of training – equals 0.177 in the production function¹⁵ and 0.121 in the wage equation. For 29 out of 33 sectors the impact of training on the marginal product of workers is higher than the impact on wages.¹⁶ Focusing on the manufacturing industries, largest productivity gains from training can be found in the Chemicals sector and Rubber and Plastic Sector.¹⁷ For the non-manufacturing sector,

¹⁵This implies the productivity premium for a trained worker β_T to be equal to .243.

¹⁶For the manufacturing and non-manufacturing sectors, respectively 14 out of 17 and 15 out of 16 sectors report a higher productivity than wage premium. Due to the relatively low number of observations in combination with the non-linear search over the parameters, the difference is however not significant for several sectors. Applying the ACF procedure, the difference is significant at the 10% level for only 10 sectors (out of 29 for which the productivity premium exceeds the wage premium). When the wage premium exceeds the productivity premium the difference is never significant. For the OLS results, the difference is significant at the 10% level for 18 out of 27 sectors for which the productivity premium is larger than the wage premium. When the wage premium exceeds the productivity premium, the difference is never significant.

¹⁷There are also large gains in the sector of Wood Products, but the training and labor coefficient are estimated imprecise.

the largest productivity gains can be found in the Agriculture, Financial Intermediation and Real Estate sectors. Figure 1 combines the estimates for the wage and productivity premia. The 45° line is plotted, such that all observations above this line represent sectors for which the impact of training on productivity is larger than the impact of training on wages. Most of the sectors are located above this line which is consistent with Acemoglu and Pischke (1998, 1999a,b).¹⁸ The correlation between the productivity and wage premium equals 0.64 and is highly significant.

Acemoglu and Pischke (1998, 1999a,b) show that firms will pay for general training when the internal wage structure is compressed, meaning that the wage function increases less steeply in general skills than the marginal product. Wage compression can be caused by a variety of labor market frictions, such as search costs and informational asymmetries leading to monopsony power. Ideally, we would like to relate our sector level estimates for the wedge between the productivity and wage premium of trained workers to a measure for monopsony power at the sector level. A positive correlation would support the view that our finding of a positive wedge between the wage and productivity premium can be best explained by a combination of general training and a compressed wage structure.

¹⁸The sectors that drop below the line, namely 14 – Mining, 21 – Paper Products and 32 – Radio, TV and Telecom. Equipment are all small sectors and we believe the low productivity premium in comparison with the wage premium is more likely due to inefficient estimates. The difference between the wage premium and productivity premium for these sectors is never statistically significant at any conventional confidence level. The fourth sector that drops below the line, 28 – Metal Products, is large however, but the productivity premium is close to the wage premium and the difference between the two is statistically insignificant.

Unfortunately, direct measures for such labor market frictions do not exist at the sector level and the estimation of monopsony power is a rather involved task, lying outside the scope of this paper.¹⁹ As an alternative, albeit far from ideal, we relate the wedge with inter-industry wage differentials. These are estimated controlling for variables mainly affecting general human capital such as education and age, but not for training. The idea is that sectors where workers are earning less than implied by their general human capital, so sectors with low inter-industry wage premia, are more monopsonistic and hence workers are less able to capture the quasi-rents of their general human capital, an argument also used by Dearden et al. (2006). Hence, we would expect a negative correlation between our estimated wedges and the inter-industry wage premia if training is general in nature. When training would be specific in nature, one would not expect the wedge to be related to the inter-industry wage premia. Using the estimates of Du Caju et al. (2010) for inter-industry wage premia in Belgium, we find that the average (median) gap between the productivity and wage premia is equal to 0.063 (0.050) in sectors for which the inter-industry wage premium is positive while the average (median) gap is equal to 0.131 (0.116) in sectors with a negative inter-industry wage premium. This tentative evidence is open to the critique that for example firm-specific training may be more prevalent in the sectors with low inter-industry wage premia. We come back to the difference between firm specific and general training in Section 6.

¹⁹For example, Manning (2003, 2011) suggests to infer the elasticity of the labor supply curve to an individual firm by estimating the wage elasticities of separations to employment and non-employment, so requiring worker-level data as well as exogenous variation in the wage rate.

5.2 Training Hours

In Table 6 we redefine the training variable as average training hours per employee and estimate equations (2) and (4) to determine the impact of training intensity on productivity and wages respectively. Again results are reported for the whole sample and manufacturing and non-manufacturing separately. We control for the possible endogeneity of training in both the production and wage equation, applying our estimation strategy described above. We estimate the productivity premium of an hour of training to be equal to 0.0076, which means that each hour of training raises the marginal product of a worker by 0.76%. The wage premium of an hour of training is estimated to be 0.44% and the difference between the wage and productivity premium is again highly significant. The results imply that the marginal product of a trained worker receiving the average amount of training hours (36.7 hours) is 27.9% higher than the marginal product of an untrained worker while its wage is only 16.1% higher, in line with the results of the previous section. Also for the manufacturing and non-manufacturing sectors separately, the productivity premium is higher than the wage premium. A summary of the results for sector specific estimates are reported in the last columns of Table 6.²⁰ Again productivity premia surpass wage premia for the majority of sectors. The correlation between the impact on productivity and on wages equals 0.49 and is highly significant.

5.3 Robustness Checks: Measurement and Estimation

The finding of substantial productivity and wage premia for trained workers where the former are larger than the latter, passes a number of robustness checks. For brevity we

²⁰In Appendix E, we graphically represent productivity and wage premia for the different sectors.

only report the wage and productivity premia as well as the test of their equality. Table 7 reports a number of robustness checks with respect to the measurement of training and the estimation method used. First, we constructed a measure for the training stock using the perpetual inventory method, i.e. $S_{it} = (1 - \delta_{it})S_{it-1} + F_{it}$ where S_{it} is the training stock of firm i in period t and F_{it} represents the training flow. Every period, the training stock depreciates at a rate of δ_{it} which consists of two components, namely the share of trained workers that leaves the firm every period and the rate at which acquired knowledge through training becomes obsolete. We approximate the first component by the observed firm level separation rate. Unfortunately we do not have information about the second component but we check the robustness of our results for different values of it (more details are provided in Appendix D). The results are reported in the first rows of Table 7. The contemporaneous impact of training on productivity and wages is estimated to be lower compared to the specifications using training flows.²¹ The difference between the wage and productivity premia remains largely significant.

Second, the Akerberg et al. (2006) methodology relies on the assumption that material input is fully flexible and that material input choices are made contemporaneously with output choices. For some sectors this assumption seems appropriate while in other sectors material orders may require substantial advance time. To identify sectors that are more likely to purchase readily available materials we combine the classification by Rauch (1999) with Supply and Use tables to distinguish between sectors using mainly homoge-

²¹This is in line with expectations. Although an increase in the stock of human capital due to training increases the contemporaneous marginal product by less, training has lingering effects and the marginal product of a trained worker remains high in the following periods.

neous and reference priced products and sectors using mainly differentiated products as inputs.²² The idea is that homogeneous and reference priced input quantities are more easily adjustable compared to input demand for differentiated products as the latter have characteristics that vary across suppliers and may even be tailored to the buyer’s needs (Besedes and Prusa, 2006). Finding new suppliers of differentiated products is more likely to involve higher search costs and to require buyer-supplier specific investments. Consequently, contract and transaction duration is likely to be longer for differentiated products compared to homogeneous and reference priced products. For example, in the international trade literature Besedes and Prusa (2006) find that differentiated products are traded longer than reference priced and homogeneous products. We estimate our main equation for the subset of industries using primarily homogeneous and reference priced inputs and results are reported in Table 7, second row. Again, training results in both positive productivity and wage premia and the former are larger than the latter.

Finally, we executed a number of additional estimation approaches. We estimate Equations (5) and (4) with Zellner’s seemingly unrelated regression (SUR) estimator.²³ Moreover, we use the average wage at a firm as a control for unobserved worker ability and finally, we include training lagged one period instead of contemporaneous training as an instrument as there could be concerns that training intensity does depend on the innovation in productivity. For example, in the case of an unexpected economic downturn firms could send their employees more easily on training since the opportunity cost of training is lower which would create a downward bias in the estimated training coefficient. For all these robustness checks, our results remained qualitatively the same as shown in the

²²More details on the the precise classification procedure are given in Appendix B.2.

²³In the SUR estimation we do not try to control for the endogeneity of inputs.

last three rows of Table 7.²⁴

5.4 Robustness Checks: Worker Heterogeneity

In Section 5.1 we made already a distinction between blue and white collar workers and assumed them to be imperfectly substitutable. In this subsection, we include other measures for worker heterogeneity, but we take the assumption that these are perfectly substitutable in order to be able to use the full data set. First we include measures for the education level of the labor force, second we control for the gender composition of the workforce.

Ideally we would be able to distinguish between high and low educated workers and observe the proportion of trained workers within each type. This would allow us to control for the education level of the workers and estimate different training premia for different types of workers. Unfortunately this information is not available, so we experimented with two different approximations to the skill level of the labor force. First, we observe the education level of every employee that leaves or enters the firm in a given year and we take the average education level of the inflow and outflow over all years as a proxy for the education level of the total labor force. We define a worker to be high-educated if he received higher or university education and low-educated if he received at most primary or secondary school education. Second, we make a distinction between blue collar workers, white collar workers and managers. We insert the shares of the different

²⁴We also estimated our main specifications using the Blundell and Bond (1998) system-GMM estimator exploiting various lag structures of the endogenous variables as instruments. While our point estimates remain robust, the Hansen test of overidentifying restrictions rejects the validity of the instruments.

types of workers in the production function and wage equation (cf. Equations 2 and 4) and apply again our methodology to control for the endogeneity of inputs.²⁵ The first part of Table 8 reports results when we control only for schooling. In the second set of results we control only for the type of contract and the third part of the table shows results of including both schooling and type of contract. More detailed results are included in Appendix G. Both the wage and productivity premium of training go down when controlling for the educational level and types of contracts. The productivity premium however is always estimated to be larger than the wage premium and the difference is statistically significant. For example, including both the type of contract and schooling level, lowers the wage premium to 9.8% and the productivity premium to 16.8%, which is more in line with results from previous studies.

Besides blue and white collar workers we observe as well the number of male and female employees. Given previous findings on productivity-wage differentials between women and men (Hellerstein and Neumark, 1999), we check the robustness of our results to the inclusion of the share of female employees. Results are reported in the final part of Table 8. By way of comparison with Table 2, it is clear that controlling for the share of male/female workers does not modify the training coefficient estimate. Appendix G reports as well results for assuming the different types of workers to be imperfectly substitutable. Again, the main results did not change.

²⁵We allow for the different contract type shares as well as the share of schooled workers to be endogenous.

6 Firm specific versus general training

In the previous sections, we have established a positive and statistically significant impact of training on productivity. Moreover the productivity premium was found to be larger than the wage premium. Note that this gap between the productivity and wage premium for trained employees can be explained equally well by perfect competition and firm specific training as by imperfect competition and general training. Each explanation however implies radically different policy implications. Which of the two theories is the best explanation for our results?

The most direct test would be looking at whether the acquired skills are transferable to other employers. For example Booth and Bryan (2005) find the wage premium of training received at previous employers is larger compared to the premium received for training at the current employer in the UK. Not only is this result a clear indication that most training is general in nature, but it also gives support to theories explaining firm provided training by labor market imperfections. Loewenstein and Spletzer (1998) find comparable results for the US. Unfortunately, we can not test directly for general versus firm specific training as we do not have employee level data about training at the current versus previous employer. However, note that our training measure represents formal training, which is more likely to be general in nature. Moreover, we attempt to infer from the turnover rates whether training is most likely to be general or specific in nature.

Under both firm specific training and general training firms are less likely to dismiss trained workers. Firm specific training is also likely to be negatively related to workers' quit rates but general training is less likely to reduce quit rates. The reason is the following. Both firm specific training and perfect competition as well as general training

and imperfect competition create a gap between the workers' wage and marginal product, making trained workers more profitable for the firm. Training costs are sunk and hence firms are less likely to fire trained workers. For workers, firm specific and general training can have a differential impact on their probability to quit the firm. Under firm specific training, acquired skills are not applicable in other firms, creating a gap between the wage of a trained worker at the current firm and the outside wage.²⁶ Consequently the probability of a voluntary quit should be lower. When training is general in nature however, it is possible that training does not have an impact on quit rates of workers. For example, when the presence of unions is the main source of wage compression, trained workers could earn the same wage at other firms leaving the quit rate unaffected at training firms. Moreover, poaching of trained workers by other firms could even increase the probability of a voluntary quit.

Our dataset allows us not only to compute general separation rates, but also to distinguish between whether these separations are dismissals initiated by the firm or quits initiated by the worker. When we regress the quit and dismissal rates on the share of trained workers lagged one and two periods, we find that dismissal rates are negatively and significantly affected by the lagged share of trained employees²⁷ as can be seen from

²⁶Note that in principle, the firm can leave the wage of trained workers unchanged after training under firm specific training as the outside option for the worker has not changed. However, Becker (1964) and Hashimoto (1981) noted that it can be optimal for both workers and firms to share benefits of training, namely under the form of higher wages but still lower than the marginal product, lowering the probability a worker quits the firm.

²⁷We control not only for firm fixed effects but include also inflows of employees both contemporaneous and lagged one period and year dummies to control for business cycles.

Table 9 . Quit rates however seem to be unaffected by the number of trained workers. The coefficient on the lagged share of trained employees is not significantly different from zero. The share of trained employees lagged two periods has even a positive and significant impact on the quit rates.²⁸ Our results are consistent with the few papers in the training literature relating job turnover to training. Lynch (1991) finds that young workers are less likely to leave the firm if they have received on-the-job training while workers that participated in off-the-job training are more likely to leave the firm. She takes this as an indication that on-the-job training is more firm specific and off-the-job training is more general. However, Parent (1999) uses the same dataset and estimates both off-the-job and on-the-job training to have a negative effect on the probability of separation. Bassanini et al. (2007) estimate the relationship between voluntary quits and training for some European countries, including Belgium, and do not find an impact of past training spells on turnover. Although not a formal proof, these results suggest that the training is most likely to be general in nature instead of firm specific and combined with our estimates of the return of training on wages and productivity give support to the theoretical work by Acemoglu and Pischke (1998) explaining training by imperfections in the labor market.

7 Conclusions

This paper used a large firm level panel data set to analyze the impact of firm provided training on both wages and productivity. We are able to measure for each firm the

²⁸When aggregating training and separation rates at the 4 digit level, there was a substantial and significantly negative correlation between the dismissal rate and the share of trained employees but not between the quit rate and share of trained employees.

amount of employees that received some kind of formal training as well as the training costs and the hours spent on training for the period 1997 to 2006. We use a control function approach to estimate production functions and wage equations at the firm level to infer productivity and wage premia of training, taking explicitly the endogeneity of training into account.

Our results indicate that the productivity increase associated with training is larger than the wage increase. More precisely, effective labor input increases by 1.7% to 3.2% in response to an increase of 10 percentage points in the fraction of workers that receive training while the average wage increases by only 1% to 1.7%. This difference between the productivity premium and the wage premium is statistically significant and robust across a wide range of specifications. We find a slightly higher impact of training in non-manufacturing compared to manufacturing sectors. Our results are robust across different specifications and definitions of the training variable. In particular, we take into account various measurement issues, estimation methods and sources of worker heterogeneity.

We provide initial evidence that the majority of training is general in nature and hence our results are consistent with recent theories such as Acemoglu and Pischke (1998, 1999a,b) which explain firm provided general training by imperfect competition in the labor market and wage compression. This finding can have important policy implications. The standard result of Becker (1964) is that if workers are not credit constrained, training investments are efficient and government intervention is unnecessary or should be directed to the credit markets. However, with imperfect labor markets and a compressed wage structure, there could be underinvestment in training from a social point of view. For example, when making their training decisions, firms do not take into account the possible externalities for future employers of trained workers (Acemoglu and Pischke 1998,

1999a,b). This opens possibilities for the government to implement training subsidies.

Tables

Table 1: Summary Statistics

	Total	Manufacturing	Non-Manufacturing
Full Sample			
Value Added ($\times 1000\text{€}$)	1,319	3,216	973
Capital Stock ($\times 1000\text{€}$)	1,043	1,985	872
Employment	21.6	40.9	18.1
Labor Costs per Worker ($\times 1000\text{€}$)	35.4	35.0	35.5
Observations	804,293	109,196	695,097
Nr. Firms	135,865	14,160	121,705
Proportion of Trained Workers	.032	.061	.026
Proportion Trained Workers in Training Firm	.503	.455	.526
Cost of Training per Worker Trained (€)	1,414	1,508	1,368
Hours Training per Worker Trained	39.1	43.1	37.2
Restricted Sample			

Value Added ($\times 1000\text{€}$)	8,288	12,727	6,343
Capital Stock ($\times 1000\text{€}$)	6,482	7,679	5,957
Employment	105.0	150.9	84.9
Labor Costs per Worker ($\times 1000\text{€}$)	47.4	43.6	48.9
Observations	91,045	27,738	63,307
Nr. Firms	15,495	4,131	11,364
Proportion of Trained Workers	.167	.222	.142
Proportion Trained Workers in Training Firm	.486	.466	.500
Cost of Training per Worker Trained (€)	1,516	1,561	1,482
Hours Training per Worker Trained	36.7	40.9	33.7

All figures refer to the unweighted average of each variable. The restricted sample refers to firms reporting material costs.

Table 2: Impact of Training on Productivity and Wages

	Total			Manufacturing			Non-Manufacturing			ACF, Lag Labor		
	OLS1	OLS2	ACF	OLS1	OLS2	ACF	OLS1	OLS2	ACF	Tot.	Man.	Non-Man.
Production Function												
Labor	.785* (.001)	.747* (.004)	.764* (.008)	.802* (.003)	.767* (.007)	.791* (.015)	.780* (.001)	.735* (.005)	.751* (.017)	.810* (.026)	.731* (.023)	.839* (.018)
Capital	.165* (.001)	.123* (.002)	.088* (.004)	.178* (.002)	.151* (.005)	.129* (.008)	.163* (.001)	.115* (.003)	.081* (.004)	.070* (.004)	.109* (.007)	.059* (.004)
Training	.460* (.008)	.315* (.010)	.243* (.010)	.403* (.015)	.300* (.016)	.215* (.017)	.181* (.016)	.461* (.008)	.301* (.012)	.223* (.015)	.192* (.016)	.234* (.014)
β_T	.586* (.010)	.422* (.014)	.318* (.014)	.502* (.019)	.391* (.021)	.272* (.022)	.591* (.013)	.410* (.017)	.342* (.018)	.276* (.015)	.262* (.025)	.278* (.019)

Wage Equation

Training (α_T)	.438*	.200*	.167*	.432*	.219*	.187*	.440*	.190*	.165*	.173*	.184*	.178*
	(.006)	(.006)	(.007)	(.009)	(.009)	(.011)	(.013)	(.007)	(.009)	(.009)	(.012)	(.013)
$\ln(K/L)$			-.015*			.017*			-.022*	-.020*	.012*	-.030*
			(.002)			(.004)			(.002)	(.002)	(.003)	(.003)
TFP			.337*			.306*			.343*	.342*	.286*	.359*
			(.006)			(.008)			(.007)	(.007)	(.010)	(.008)

Wald Test $\beta_T = \alpha_T$

χ^2_1	341.1	333.6	127.1	32.1	89.4	14.1	245.0	231.6	113.0	35.7	18.7	17.4
$p - value$.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Obs	804,293	73,930	73,930	123,834	23,345	23,345	677,764	50,585	50,585	63,134	20,041	43,093
Clust	135,865	13,757	13,757	18,422	3,878	3,878	117,021	9,879	9,879	12,520	3,587	8,933

Estimates for production function and wage equation, all sectors pooled together. *OLS1* refers to OLS estimation of both the production function on the full sample. *OLS2* refers to the same estimation equation, but now for the subsample of firms reporting material costs. In *ACF* we control for the endogeneity of inputs and training. The last three columns report results using lags of labor instead of current labor as instrument. Standard errors are computed using a block bootstrap procedure with 500 replications and are robust against heteroskedasticity and intra-group correlation.

Table 3: Blue versus White Collar Workers, Imperfect Substitution

	OLS		ACF	
	Prod.	Wage	Prod.	Wage
Capital	.163*		.113*	
	(.005)		(.005)	
Blue Collar	.295*		.275*	
	(.006)		(.049)	
White Collar	.448*		.452*	
	(.011)		(.011)	
Training β_T or α_T	.297*	.163*	.228*	.130*
	(.016)	(.004)	(.022)	(.008)
Nr. Obs.	46,052		46,052	
Nr. Clust.	8,753		8,753	
Test for $\beta_T = \alpha_T$				
Chi^2	78.7		16.0	
$p - value$.000		.000	

Results ACF method blue collar workers lagged one period as instrument. Standard errors are computed using a block bootstrap procedure with 500 replications and are robust against heteroskedasticity and intra-group correlation. * Significant at 5%.

Table 4: Results Manufacturing Sectors

	OLS			ACF			Nr. Obs	
	β_T	α_T	$\beta_T = \alpha_T$	β_T	α_T	$\beta_T = \alpha_T$	Obs	Clust.
15 Food Products	.209* (.042)	.155* (.020)	2.38	.178* (.047)	.120* (.025)	1.41	3,571	600
17 Textile Products	.423* (.060)	.171* (.031)	19.9	.313* (.069)	.121* (.039)	5.84	1,729	298
18 Wearing Apparel	.152 (.283)	.123 (.216)	.01	.071 (.175)	.031 (.191)	.02	364	66
20 Wood Products	.986* (.222)	.185* (.044)	14.1	.378 [†] (.226)	.110* (.055)	1.52	588	110
21 Paper Products	.174* (.089)	.179* (.035)	0.00	.055 (.085)	.138* (.040)	1.04	685	98
22 Publishing	.191* (.089)	.074* (.035)	2.93	.190* (.085)	.078* (.040)	3.42	1,764	319

24 Chemical Products	(.075)	(.027)	.087 [†]	(.059)	(.029)	.064 [†]		
	.482*	.312*	5.95	.374*	.242*	3.00	2,134	331
25 Rubber and Plastic	(.079)	(.033)	.015*	(.075)	(.042)	.083 [†]		
	.433*	.239*	10.66	.319*	.211*	.62	1,477	227
26 Mineral Products	(.064)	(.029)	.001*	(.113)	(.039)	.423		
	.274*	.155*	4.57	.224*	.144*	1.85	1,917	311
27 Basic Metals	(.061)	(.024)	.033*	(.062)	(.027)	.174		
	.131	.146*	.05	.113	.109*	.00	995	146
28 Metal Products	(.085)	(.049)	.819	(.088)	(.049)	.967		
	.231*	.195*	.57	.138*	.157*	.14	2,779	478
29 Machinery	(.050)	(.031)	.451	(.056)	(.032)	.710		
	.349*	.237*	4.07	.248*	.198	0.34	1,681	283
31 Electrical Machinery	(.063)	(.028)	.044*	(.078)	(.038)	.561		
	.294*	.218*	.42	.296*	.171*	.68	665	104

	(.123)	(.046)	.517	(.151)	(.055)	.411		
32 Radio, TV and Telecom	.225 [†]	.287*	.35	.169	.248*	.29	302	54
	(.123)	(.061)	.556	(.170)	(.079)	.591		
33 Medical and Prec. Eq.	.315*	.104*	3.01	.230	.073	1.12	346	64
	(.130)	(.045)	.083 [†]	(.169)	(.056)	.289		
34 Motor Vehicles	.060	.123*	.089	.113	.107*	.01	752	121
	(.083)	(.043)	.344	(.076)	(.046)	.930		
36 Furniture	.130	.177*	.25	.252*	.142*	.75	1,101	181
	(.094)	(.049)	.618	(.122)	(.056)	.385		

Productivity and wage premium for non-manufacturing sectors. OLS refers to OLS estimation of the production function and wage equation without control for TFP on the subsample of firms that report material inputs. ACF refers to the results of the production function and wage equation controlling for the endogeneity of inputs. Standard errors are computed using a block bootstrap procedure with 500 replications and are robust against heteroskedasticity and intra-group correlation. * Significant at 5 %, † significant at 10 %.

Table 5: Results Non-Manufacturing Sectors

	OLS			ACF		Nr. Obs	
	β_T	α_T	$\beta_T = \alpha_T$	β_T	α_T	$\beta_T = \alpha_T$	Obs Clust.
1 Agriculture	.280 (.247)	.125 (.090)	.600 .438	.524* (.244)	.145† (.076)	2.65 .010*	459 96
14 Mining	.387* (.182)	.091 (.066)	3.25 .071†	.016 (.172)	.054 (.074)	.06 .807	345 60
37 Recycling	.324† (.198)	.267* (.094)	.10 .756	.328 (.261)	.254* (.094)	.09 .762	364 79
45 Construction	.238* (.029)	.150* (.019)	9.75 .002*	.159* (.031)	.134* (.020)	.60 .440	5,521 939
50 Sales Motor Vehicles	.368* (.053)	.185* (.024)	18.1 .000*	.223* (.044)	.138* (.028)	3.39 .066†	3,970 745

51 Wholesale Trade	.473*	.191*	126.1	.418*	.181*	82.6	21,380	4,017
	(.029)	(.013)	.000*	(.030)	(.014)	.000*		
52 Retail Trade	.160*	.077*	5.81	.184*	.060*	3.74	4,104	869
	(.042)	(.022)	.016*	(.062)	(.025)	.053 [†]		
55 Hotels and Restaurants	.130*	.122*	.03	.120*	.098*	.12	877	164
	(.052)	(.029)	.863	(.059)	(.034)	.728		
60 Land Transport	.154 [†]	.094*	.97	.110 [†]	.061*	.69	2,617	455
	(.083)	(.029)	.325	(.063)	(.028)	.401		
63 Transport Activities	.481*	.121*	19.7	.326*	.100*	10.2	1,892	411
	(.090)	(.032)	.000*	(.081)	(.034)	.000		
64 Post and Telecom.	.376	-.014	2.78	.379	-.082	3.33	337	86
	(.257)	(.105)	.096 [†]	(.24)	(.110)	.068 [†]		
65 Financial Intermed.	.720*	.025	3.86	.591 [†]	-.021	3.22	315	79
	(.335)	(.105)	.050 [†]	(.310)	(.118)	.073 [†]		

70 Real Estate	.639*	.303*	4.51	.482*	.265*	1.10	1,360	275
	(.179)	(.063)		(.239)	(.072)	.294		
71 Renting of Machinery	.205	.129*	.12	.259	.084	1.17	487	96
	(.238)	(.056)	.731	(.165)	(.050)	.280		
72 Computer Services	-.035	-.024	.09	.002	-.005	.03	1,587	393
	(.046)	(.031)	.770	(.052)	(.036)	.866		
74 Business Services	.252*	.143*	9.80	.259*	.137*	6.37	4,313	975
	(.039)	(.022)	.002*	(.053)	(.025)	.011*		

Productivity and wage premium for non-manufacturing sectors. OLS refers to OLS estimation of the production function and wage equation without control for TFP on the subsample of firms that report material inputs. ACF refers to the results of the production function and wage equation controlling for the endogeneity of inputs. Standard errors are computed using a block bootstrap procedure with 500 replications and are robust against heteroskedasticity and intra-group correlation. * Significant at 5 %, † significant at 10 %.

Table 6: Training as Average Training Hours per Worker

	Total	Manuf.	Non-Manuf.	Each Sector Separat.	
Production Function					
Labor	.770*	.794*	.756*	β_T	
	(.008)	(.016)	(.009)	Min	-.0015
Capital	.089*	.131*	.082*	Max	.0168
	(.004)	(.001)	(.004)	Av.	.0059
Training Hours	.0058*	.0047*	.0065*		
	(.0003)	(.0004)	(.0004)		
β_T	.0076*	.0058*	.0087*		
	(.0004)	(.0005)	(.0005)		
Wage Equation					
Training Hours(α_T)	.0044*	.0044*	.0046*	α_T	
	(.0002)	(.0003)	(.0003)	Min	-.0008
ln(K/L)	-.015*	.018*	-.022*	Max	.0099
	(.002)	(.004)	(.002)	Av.	.0032
TFP	.340*	.311*	.346*		
	(.006)	(.009)	(.007)		
Wald Test $\beta_T = \alpha_T$					
χ^2_1	86.0	7.1	69.2		
$p - value$.000	.008	.000		

ACF procedure to estimate wage equation and production function. Standard errors are computed using a block bootstrap procedure with 500 replications and are robust against heteroskedasticity and intra-group correlation.* Significant at 5%.

Table 7: Results Further Robustness Checks

	Prod.	Wage	Test for Equality $\beta_T = \alpha_T$	
	β_T	α_T	χ^2_1	$p - value$
Training Stock	.181*	.090*	136.0	.000
	(.009)	(.003)		
Fully Flexible Materials	.263*	.180*	7.58	.006
	(.029)	(.016)		
Wage as Control	.250*	.168*	261.4	.000
	(.014)	(.006)		
SUR Model	.391*	.208*	593.8	.000
	(.008)	(.005)		
Lagged Training	.224*	.147*	30.3	.000
	(.015)	(.007)		

Different robustness checks. First, training intensity is measured by the training stock. Second, results for the subsample of sectors using reference priced or homogenous goods are reported. Third, we use wage as a control variable for labor quality in the production function. Fourth, we apply Zellner's SUR estimator and fifth, we use training lagged one period as instrument. * Significant at 5%.

Table 8: Worker Heterogeneity, Perfect Substitution

	OLS		ACF	
	Prod.	Wage	Prod.	Wage
Schooling (β_T or λ_T)	.255*	.112*	.189*	.094*
	(.016)	(.005)	(.012)	(.013)
Wald Test $\beta_T = \alpha_T$				
χ_1^2 (p - value)	161.4	(.000)	18.1	(.000)
Type of Contract (β_T or α_T)	.273*	.155*	.222*	.139*
	(.012)	(.005)	(.013)	(.006)
Wald Test $\beta_T = \alpha_T$				
χ_1^2 (p - value)	123.4	(.000)	45.3	(.000)
Type of Contract and Schooling (β_T or α_T)	.211*	.112*	.168*	.098*
	(.011)	(.005)	(.012)	(.008)
Wald Test $\beta_T = \alpha_T$				
χ_1^2 (p - value)	99.7	(.000)	24.6	(.000)
Female/Male Employees (β_T or α_T)	.417*	.199*	.301*	.164*
	(.014)	(.006)	(.013)	(.007)
Wald Test $\beta_T = \alpha_T$				
χ_1^2 (p - value)	360.7	(.000)	129.6	(.000)

Results of controlling for different types of worker heterogeneity. Full sample pooled. Standard errors are computed using a block bootstrap procedure with 500 replications and are robust against heteroskedasticity and intra-group correlation. * Significant at 5%.

Table 9: Separation rates and training

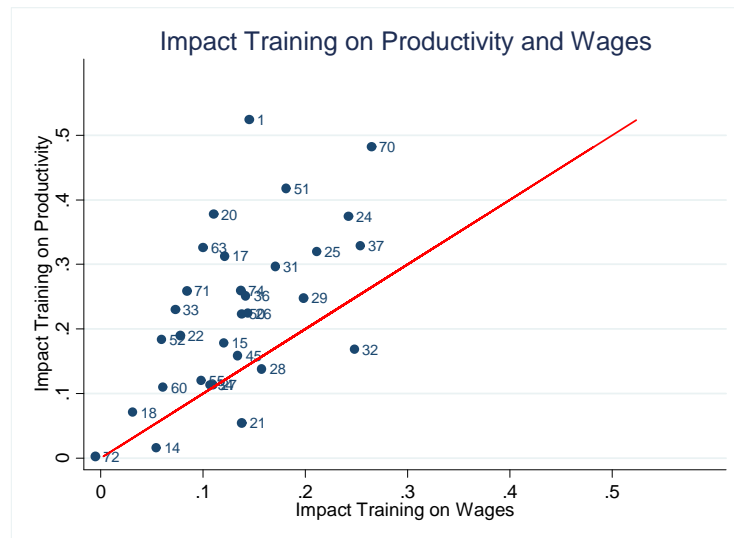
	FE one lag		FE two lags	
	Dismissals	Quits	Dismissals	Quits
Train. Share _{t-1}	-.00254 [†]	-.00132	-.0028 [†]	-.0023
	(.0015)	(.0024)	(.0015)	(.00241)
Train. Share _{t-2}			.00174	.00627*
			(.00149)	(.00237)
Nr. Obs	76,359	76,359	76,340	76,340

Firm and year fixed effects as well as the inflow of employees both contemporaneous and lagged one period included. [†]

Significant at 10%. * Significant at 5%.

Figures

Figure 1: Impact Training on Productivity and Wages



References

- Acemoglu, Daron, and Jörn-Steffen Pischke, “Why Do Firms Train? Theory and Evidence,” *Quarterly Journal of Economics* 113(1) (1998), 79-119.
- Acemoglu, Daron, and Jörn-Steffen Pischke, “The Structure of Wages and Investment in General Training,” *Journal of Political Economy* 107(3) (1999a), 539-572.
- Acemoglu, Daron, and Jörn-Steffen Pischke, “Beyond Becker: Training in Imperfect Labour Markets,” *The Economic Journal* 109(453) (1999b), 112-142.
- Akerberg, Daniel, Richard Caves, and Garth Frazer, “Structural Identification of Production Functions,” mimeograph, UCLA (2006).
- Altonji, Joseph G., and James R. Spletzer, “Worker Characteristics, Job Characteristics, and the Receipt of On-the-Job Training,” *Industrial and Labor Relations Review* 45(1) (1991), 58-79.
- Bartel, Ann P. “Training, Wage Growth, and Job Performance: Evidence from a Company Database,” *Journal of Labor Economics* 13(3) (1995), 401-425.
- Bassanini, Andrea, Alison L. Booth, Giorgio Brunello, Maria De Paola and Edwin Leuven, “Workplace Training in Europe,” in Giorgio Brunello, Pietro Garibaldi, and Etienne Wasmer (eds.), *Education and Training in Europe* (Oxford: Oxford University Press, 2007).
- Becker, Gary S., *Human Capital* (Chicago: The University of Chicago Press, 1964).
- Besedeš, Tibor, and Thomas J. Prusa, “Product Differentiation and Duration of US Import Trade,” *Journal of International Economics* 70(2) (2006), 339-358.

Black, Sandra E., and Lisa M. Lynch,. “How To Compete: The Impact Of Workplace Practices and Information Technology on Productivity,” *The Review of Economics and Statistics* 83(3) (2001), 434-445.

Blundell, Richard, and Stephen Bond, “Initial Conditions and Moment Restrictions in Dynamic Panel Data Models,” *Journal of Econometrics* 87(1) (1998), 115-143.

Booth, Alison L., “Job-Related Training: Who Receives it and What Is it Worth?,” *Oxford Bulletin of Economics and Statistics* 53(3) (1991), 281-294.

Booth, Alison L., Marco Francesconi, and Gylfi Zoega, “Unions, Work- Related Training, and Wages: Evidence for British Men,” *Industrial & Labor Relations Review* 57(1) (2003), 68-91.

Booth, Alison L., and Mark L. Bryan, “Testing Some Predictions of Human Capital Theory: New Training Evidence from Britain,” *The Review of Economics and Statistics* 87(2) (2005), 391-394.

Dearden, Lorraine, Howard Reed, and John Van Reenen, “The Impact of Training on Productivity and Wages: Evidence from British Panel Data,” *Oxford Bulletin of Economics and Statistics* 68(4) (2006), 397-421.

Du Caju, Philip, Ana Lamo, Steven Poelhekke, Gabor Kátay, and Daphne Nicolitsas, “Inter-Industry Wage Differentials in EU Countries: What do Cross-Country Time Varying Data Add to the Picture?,” *Journal of the European Economic Association* 8(2-3) (2010), 478-486.

Frazer, Garth “Linking Firms and Workers: Heterogeneous Labor and Returns to Education,” mimeograph, Yale University (2001).

Frazis, Harley, and Mark A. Loewenstein, "Reexamining the Returns to Training: Functional Form, Magnitude, and Interpretation," *The Journal of Human Resources* 40(2) (2005), 453-476.

Hashimoto, Masanori, "Firm-Specific Human Capital as a Shared Investment," *American Economic Review* 71(3) (1981), 475-482.

Hellerstein, Judith K., and David Neumark, "Sex, Wages, and Productivity: An Empirical Analysis of Israeli Firm-Level Data," *International Economic Review* 40(1) (1999), 95-123.

Hellerstein, Judith K., David Neumark, and Kenneth R. Troske, "Wages, Productivity, and Worker Characteristics: Evidence from Plant Level Production Functions and Wage Equations," *Journal of Labor Economics* 17(3) (1999), 409-446.

Loewenstein, Mark A., and James R. Spletzer, "General and Specific Training: Evidence and Implications," *The Journal of Human Resources* 34(4) (1999), 710-733.

Lucas, Robert E., "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22(1) (1988), 3-42.

Lynch, Lisa M., "Private-Sector Training and the Earnings of Young Workers," *American Economic Review* 82(1) (1992), 299-312.

Manning, Alan, *Monopsony in Motion* (Princeton, NJ: Princeton University Press, 2003).

Manning, Alan, "Imperfect Competition in the Labor Market," in David Card and Orley Ashenfelter (eds.), *Handbook of Labor Economics*, vol. 4B (Amsterdam: North-Holland, 2011).

Marschak, Jacob, and William H. Jr. Andrews, “Random Simultaneous Equations and the Theory of Production,” *Econometrica* 12(3/4) (1944), 143-205.

Mincer, Jacob, *Schooling, Experience, and Earnings* (New York: Colombia Univ. Press, 1974)

OECD, *OECD Economic Surveys: Belgium* (Paris: OECD, 2007).

Parent, Daniel, “Wages and Mobility: The Impact of Employer Provided Training,” *Journal of Labor Economics* 17(2) (1999), 298-317.

Pischke, Jörn-Steffen, “Continuous Training in Germany,” *Journal of Population Economics* 14(3) (2001), 523-548.

Rauch, James E., “Networks versus Markets in International Trade,” *Journal of International Economics* 48(1) (1999), 7-35.

Van Biesebroeck, Johannes, “Wages Equal Productivity. Fact or Fiction? Evidence from Sub Saharan Africa,” *World Development* 39(8) (2011), 1333-1346.

Zwick, Thomas, “The Impact of Training Intensity on Establishment Productivity,” *Industrial Relations: A Journal of Economy and Society* 45(1) (2006), 26-46.